

AD-A285 922



ARMY RESEARCH LABORATORY



# A Multivariate Multisample Rank Test for Stochastic Simulation Validation

Ann E. M. Brodeen  
Malcolm S. Taylor

ARL-TR-592

October 1994



94-34309



DTIC QUALITY INSPECTED 5

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

94 11 3 06 6'

## **NOTICES**

**Destroy this report when it is no longer needed. DO NOT return it to the originator.**

**Additional copies of this report may be obtained from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.**

**The findings of this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.**

**The use of trade names or manufacturers' names in this report does not constitute indorsement of any commercial product.**

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE October 1994	3. REPORT TYPE AND DATES COVERED Final, 1 October 1993-31 May 1994		
4. TITLE AND SUBTITLE A Multivariate Multisample Rank Test for Stochastic Simulation Validation		5. FUNDING NUMBERS WO: 4B592-402-34-3003		
6. AUTHOR(S) Ann E. M. Brodeen and Malcolm S. Taylor				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: AMSRL-CI-CC Aberdeen Proving Ground, MD 21005-5067		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: AMSRL-OP-AP-L Aberdeen Proving Ground, MD 21005-5066		10. SPONSORING/MONITORING AGENCY REPORT NUMBER ARL-TR-592		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words)  Nonparametric multivariate statistical methods address a broad class of problems in data analysis in which the assumption of normality is not feasible, and where the data occur naturally as n-tuples (vectors) rather than scalar values. This is the data structure that is most common in the engineering sciences and, coincidentally, the least tractable. A computer-intensive approach to the analysis of these data, usually referred to as randomization or permutation procedures, will be the specific focus of this work. Tests based on permutations of observations are nonparametric tests. This study considers a multivariate extension of the well-known Kruskal-Wallis rank sum test as a method for hypothesis testing, a technique commonly employed for simulation validation. The test statistic investigated is a nonparametric analogue of the classical Hotelling $T^2$ used for the normal theory model. This undertaking is part of a broader based Army research program, the goal of which is to improve the ability of communications networks to deliver critical information on the battlefield when and where it is needed, despite a rapidly changing and often hostile environment. It will also support the ongoing effort to formalize the validation process for network simulation that, in turn, provides the groundwork for exploring alternatives and testing hypotheses throughout the research program. This formalization of the validation process can be readily transmitted to other organizations that rely on network simulations for their analyses.				
14. SUBJECT TERMS multivariate analysis, permutation tests, simulation validation, command and control, protocol performance, networks			15. NUMBER OF PAGES 32	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

INTENTIONALLY LEFT BLANK.

## TABLE OF CONTENTS

	<u>Page</u>
LIST OF FIGURES .....	v
LIST OF TABLES .....	v
1. INTRODUCTION .....	1
1.1 Limited Bandwidth Tactical Networks .....	1
1.2 Network Effectiveness .....	1
1.3 Experimentation vs. Simulation .....	1
1.4 The Verification, Validation, and Accreditation (VV&A) Process .....	1
1.5 Current Research .....	2
2. PERMUTATION TESTS .....	2
2.1 Conditional Nonparametric Hypothesis Tests .....	2
2.2 General Setting for Rank Statistics .....	3
2.3 Theorem 2.3 .....	3
2.4 Theorem 2.4 .....	4
2.5 Permutation Principle .....	5
2.6 Two-Sample Univariate Location Problem .....	5
2.7 Multivariate Extension .....	6
2.8 Theoretical Development .....	6
2.9 Matrix Representation .....	7
2.10 Data Analysis .....	8
2.11 Rank Matrix Representation .....	8
2.12 Multivariate Analogue of the Permutation Principle .....	8
2.13 Average Rank Scores .....	9
2.14 Rank Order Test Statistic .....	9
2.15 The Discriminant .....	10
2.16 The Covariance .....	11
2.17 The Quadratic Form .....	11
2.18 Application .....	11
3. TACTICAL NETWORK EXPERIMENT AND SIMULATION DEVELOPMENT .....	11
3.1 Background .....	11
3.2 Experimental Design Factors of Interest .....	11
3.3 Experimental and Simulation Configurations .....	12

	<u>Page</u>
4. APPLICATION TO NETWORK SIMULATION VALIDATION .....	17
4.1 Foreword .....	17
4.2 Reformulation of the Location Problem .....	17
4.3 MOP .....	18
4.4 Case Selection .....	18
4.5 Testing the Null Hypothesis .....	19
4.6 Remarks .....	20
4.7 Chung and Fraser Test Statistic .....	20
5. PROJECTION OF CURRENT RESEARCH .....	21
6. SUMMARY .....	24
REFERENCES .....	25
BIBLIOGRAPHY .....	27
DISTRIBUTION LIST .....	29

<b>Accession For</b>	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability	
Dist	Special
A-1	

## LIST OF FIGURES

1.	Hardware configuration .....	13
2.	Simulation configuration .....	14
3.	The node model .....	14
4.	The server model .....	15
5.	Media access control process model .....	16
6.	Marginals for a representative data set - throughput vs. utilization .....	22
7.	Marginals for a representative data set - delay vs. utilization .....	22
8.	Marginals for a representative data set - throughput vs. delay .....	23

## LIST OF TABLES

1.	Permutation Values of $\bar{y} - \bar{x}$ .....	7
2.	Multivariate Multisample Rank Sum Test Results .....	20
3.	Chung and Fraser Test Statistic Results .....	21

INTENTIONALLY LEFT BLANK.

## 1. INTRODUCTION

1.1 Limited Bandwidth Tactical Networks. The purpose of a network is to serve as a carrier of information from one point to another. On a limited bandwidth tactical network, the number of nodes and the amount of information to pass can be large, especially during peak battle periods. The effective distribution of information can enhance the decision process on the battlefield, while the impact of making decisions from old information can be catastrophic (Brodeen, Kaste, and Broome 1992).

1.2 Network Effectiveness. To measure a network's effectiveness, one must determine whether the messages the network services arrive at their destination correctly and in time to be useful. The amount of correctly passed information is referred to as "network throughput," and the amount of time required to pass that information as "network delay." There are a number of parameters that can impact throughput and delay; for example, the number of messages to be transmitted, the size of the messages, the number of nodes on the network, the communications protocol, and the communications hardware. If the interaction of these network parameters is understood, the network's effectiveness can be optimized.

1.3 Experimentation vs. Simulation. One way to examine the interaction of network parameters is through simulation. But communications protocols are often too complex to model precisely. The simulations often take required input, such as the probability two or more messages will collide, the expected delay in message transmission, or the arrival rate of messages at a given node, and extrapolate those estimates to a large scenario of multiple nodes. These drastic assumptions, usually made to simplify the simulation, may actually result in an unrealistic representation of the protocol. Controlled experimentation with the actual communications protocol on the intended hardware offers much insight into the behavior of the protocol under various conditions, facilitating the modeling and simulation efforts (Brodeen, Kaste, and Broome 1992).

1.4 The Verification, Validation, and Accreditation (VV&A) Process. Recent events have caused the defense community (e.g., various defense and service science boards, the General Accounting Office [GAO], the Defense Modeling and Simulation Office [DMSO], etc.) to refocus considerable attention on the VV&A process of the models and simulations it uses. A forthcoming Department of Defense (DOD) directive on modeling and simulation will require each military service to establish VV&A policies, guidelines, and procedures. The research outlined in this paper presents an enhancement to the formal results validation procedure. Results validation will hereafter refer to the formal documented review

process that compares responses of a model and/or simulation with known or expected behavior from the subject or system it represents to ascertain that the model/simulation responses are sufficiently accurate for intended uses. A variety of methods may be employed in results validation: comparison with expert expectation (i.e., high face validation), actual test data, results from other models, or historical data (Sargent 1992).

- **Objectives and Challenges.** Experimentation with a simulation is only a surrogate for actually being able to experiment with an existing or proposed system. A reasonable goal of validation is to ensure that a simulation is developed that can actually be used by a decision maker to make the same decision that would have been made if it were feasible and cost-effective to experiment with the actual system. Validation should enhance the confidence placed in the results produced by the simulation. The challenge is to develop a validation process that is at the same time feasible yet more effective, and can be applied to both existing simulations as well as newly developed ones.

1.5. Current Research. Simulation and modeling are widely accepted means of analyzing real-world systems that are too complex to model analytically. Most communications networks fall into this category. But model credibility suffers when a continuing verification and validation program is not undertaken, thereby diluting the value of analyses the models support. It is not uncommon within a military organization to find several groups each developing a network simulation that performs essentially the same tasks; the differences usually lie in the model assumptions and/or definitions of simulation responses. An independent evaluator is called upon to assess the performance of several simulations against limited empirical data. The product of this research will be to formalize a multivariate multisample rank sum test that will enhance long-term efforts to standardize the process of building, verifying, and validating command, control, and communications (C3) simulations for flexibly addressing issues related to low-level information distribution on the battlefield. This research will also serve to strengthen the link between experimentation and simulation, both of which should be utilized in evaluating communications protocols' measures of performance (MOP).

## 2. PERMUTATION TESTS

2.1 Conditional Nonparametric Hypothesis Tests. In this section, we consider the construction of nonparametric (distribution-free) hypothesis tests whose critical regions are determined from information gained from observed data. The critical region is thus conditional, since it can be created only after the

data have been observed. Nonetheless, the test procedure has overall significance level  $\alpha$  because the critical region is constructed to assure the conditional probability of rejecting a valid null hypothesis  $H_0$  remains  $\alpha$ . Conditional hypothesis tests are discussed at several levels of theoretical intensity, ranging from Conover (1971), Noreen (1989), Randles and Wolfe (1979), and Edgington (1987) to Puri and Sen (1993). Our ultimate interest lies in hypothesis testing in a multivariate multisample framework; but to fix ideas and, to some extent, notation, we begin with consideration of a two-sample univariate location problem.

**2.2 General Setting for Rank Statistics.** Let  $X_1, \dots, X_m$  and  $Y_1, \dots, Y_n$  be independent random samples from continuous distributions with cumulative distribution functions (c.d.f.)  $F(x)$  and  $F(x - \delta)$ , respectively, where  $-\infty < \delta < \infty$ , and define

$$\begin{aligned} Z_i &= X_i, & i &= 1, \dots, m \\ &= Y_{i-m}, & i &= m+1, \dots, N, \end{aligned} \quad (1)$$

with  $N = m + n$ . Let  $Z_{(1)} \leq \dots \leq Z_{(N)}$  denote the combined sample order statistics and

$$Z_{(\cdot)} = (Z_{(1)}, \dots, Z_{(N)}) \quad (2)$$

the vector of order statistics. If the distributions of the random variables  $X$  and  $Y$  are identical (i.e.,  $\delta = 0$ ), then every arrangement of the  $X$ 's and  $Y$ 's in the ordered combined sample should be equally likely. This is the basic principle underlying many nonparametric procedures based on ranks, and is established formally as Theorem 2.3.

**2.3 Theorem 2.3.** Let  $Z_1, \dots, Z_N$  be a random sample from a continuous distribution, and let  $R = (R_1, \dots, R_N)$  be the corresponding vector of ranks (i.e.,  $Z_i = Z_{(R_i)}$ ,  $i = 1, \dots, N$ ). If  $R$  is the permutation group of the integers  $1, \dots, N$ , then  $R$  is distributed uniformly over  $R$  (Randles and Wolfe 1979).

• **Proof.** The permutation group  $R$  has  $N!$  elements. It will suffice to show that  $R$  assumes each of the permutations of  $(1, \dots, N)$  with probability  $1/N!$ . Let  $r = (r_1, \dots, r_N) \in R$  be an arbitrary element of  $R$ . Then

$$\begin{aligned} P(R = r) &= P\left[(Z_1, \dots, Z_N) = (Z_{(r_1)}, \dots, Z_{(r_N)})\right] \\ &= P(Z_{d_1} < \dots < Z_{d_N}), \end{aligned} \quad (3)$$

where  $d_i$  is the location of the number  $i$  in the permutation  $r$ , for  $i = 1, \dots, N$ . But  $Z_1, \dots, Z_N$  are independent and identically distributed (i.i.d.) random variables, therefore,

$$P(R = r) = P(Z_{d_1} < \dots < Z_{d_N}) = P(Z_1 < \dots < Z_N) = P(R = r_0) \quad (4)$$

where  $r_0 = (1, \dots, N)$ . But  $r$  is arbitrary, and the cardinality of  $R$  is  $N!$ , hence,

$$P(R = r) = 1/N! \quad (5)$$

This completes the proof.

**2.4. Theorem 2.4.** Let  $Z_1, \dots, Z_N$  be i.i.d. continuous random variables, and let  $R = (R_1, \dots, R_N)$  denote the rank vector of these observations; that is,  $R_i$  is the rank of  $Z_i$  among  $Z_1, \dots, Z_N$ . Let  $Z_{(1)} < \dots < Z_{(N)}$  be the order statistics of  $Z_1, \dots, Z_N$ .  $R$  and  $Z_{(1)} < \dots < Z_{(N)}$  are independent (Randles and Wolfe 1979).

• **Proof.** It will suffice to show that the conditional distribution of  $Z_{(1)} < \dots < Z_{(N)}$ , given  $R = r$ , is equal to the marginal distribution of  $Z_{(1)} < \dots < Z_{(N)}$ , for arbitrary  $r \in R$ . Consider  $r = r^* = (1, \dots, N)$ . Then  $Z_{(1)} = Z_1, \dots, Z_{(N)} = Z_N$  and

$$\begin{aligned}
g(z_{(1)}, \dots, z_{(N)} | R = r^*) &= \prod_{i=1}^N f(z_{(i)}) / P[R = (1, \dots, N)] \\
&= N! \prod_{i=1}^N f(z_{(i)}),
\end{aligned} \tag{6}$$

which is the joint unconditional distribution of  $Z_{(1)}, \dots, Z_{(N)}$ . This completes the proof.

**2.5 Permutation Principle.** With the aid of Theorems 2.3 and 2.4, the expression

$$P_0(Z_1 = z_{(r_1)}, \dots, Z_N = z_{(r_N)} | Z_{(\cdot)} = z_{(\cdot)}) = 1/N! \quad \forall r \in R \tag{7}$$

is established. This equation is the mathematical statement of the permutation principle, and it provides the basis for construction of conditional distribution-free tests of hypotheses. Conditioning on the order statistics vector  $Z_{(\cdot)} = z_{(\cdot)}$  reflects that the data have been observed. Under the assertion of identical distributions ( $H_0: \delta = 0$ ), the population labels are suppressed, and every arrangement of the data is equally likely. The transformation from the permutation principle (equation 7) to the mechanics of hypothesis test construction is best conveyed by example.

**2.6 Two-Sample Univariate Location Problem.** Taylor (1992) presents two sets of measurements made on spin rates of long-rod penetrators corresponding to two distinct fin configurations, where  $x_1 = 97.5$ ,  $x_2 = 122.2$ ,  $x_3 = 108.2$ , and  $y_1 = 78.1$ ,  $y_2 = 76.7$ ,  $y_3 = 88.5$  are the observed values of random samples of sizes  $m = n = 3$  from continuous distributions with c.d.f.'s  $F(x)$  and  $F(x - \delta)$ , respectively. The observed order statistics of the combined sample are then  $z_{(1)} = 76.7$ ,  $z_{(2)} = 78.1$ ,  $z_{(3)} = 88.5$ ,  $z_{(4)} = 97.5$ ,  $z_{(5)} = 108.2$ , and  $z_{(6)} = 122.2$ . Since there is no constraint as to which of the two fin configurations might provide the larger mean, a two-tailed test is appropriate. To construct a conditional test that is a distribution-free permutation test of  $H_0: \delta = 0$  against the alternative  $H_1: \delta \neq 0$ , we desire a statistic,  $S(X_1, X_2, X_3; Y_1, Y_2, Y_3)$ , that is a measure of  $\delta$ . In this example, we select  $S(X_1, X_2, X_3; Y_1, Y_2, Y_3) = \bar{Y} - \bar{X}$  with  $\bar{X} = \frac{1}{3} \sum_{i=1}^3 X_i$  and  $\bar{Y} = \frac{1}{3} \sum_{j=1}^3 Y_j$ . Next, we compute the  $C = \binom{6}{3} = 20$  possible values of  $\bar{Y} - \bar{X}$  corresponding to the ways in which we can assign three of the ordered values  $z_{(1)}, \dots, z_{(6)}$  to be designated as  $x$ 's. The resulting values of  $S$ , denoted by  $s_1, \dots, s_{20}$ , are given in Table 1. The

associated conditional null distribution of  $\bar{Y} - \bar{X}$ , given  $Z_{(\cdot)} = z_{(\cdot)}$ , then assigns probability  $1/20$  to each of the  $s_i$  values in Table 1. Since both small and large values of  $\bar{Y} - \bar{X}$  are indicative of the alternative  $H_1: \delta \neq 0$ , the critical regions for the corresponding permutation test would contain an appropriate number of the largest  $|s_i|$  values. The proportion of the data permutations with as large a value of  $|\bar{y} - \bar{x}|$  as 28.200 is chosen. Thus, an  $\alpha = .10$  ( $= 2/20$ ) level critical region would be  $C_{.10} = \{28.200, -28.200\}$ . This is the smallest level at which this permutation test would reject  $H_0: \delta = 0$  in favor of  $H_1: \delta \neq 0$ . The data for this example corresponds to  $s_{20} = -28.200$ .

**2.7 Multivariate Extension.** The permutation principle, which has thus far been restricted to a univariate two-sample setting, can be extended and applied to a wide variety of statistical problems. The construction of conditional tests finds application in fundamental considerations of multivariate analysis where counting and ranking techniques do not lend themselves effectively to small sample situations. Puri and Sen (1993) provide a rigorous treatment of the use of conditional tests in dealing with problems in multivariate data analysis. The approach in the following sections corresponds in the main to their development.

**2.8 Theoretical Development.** Let

$$X_j^k = (X_{1j}^k, \dots, X_{pj}^k)', \quad j = 1, \dots, n_k, \quad k = 1, \dots, c \quad (8)$$

be independent  $p$ -dimensional random variables from  $c$  continuous distributions with the c.d.f. of  $X^k$  denoted by  $F_k(x)$ ,  $k = 1, \dots, c$ . The data structure is that of a multivariate multisample ( $2 \leq c$ ) location problem; i.e.,

$$F_k(x) = F(x - \delta_k), \quad k = 1, \dots, c, \quad (9)$$

and the interest is in testing  $H_0: \delta_1 = \dots = \delta_c$  against the alternative  $\delta_r \neq \delta_s$  for some  $r \neq s$ .

Table 1. Permutation Values of  $\bar{y} - \bar{x}$

Values Assigned as $x_i$ 's	Values of $\bar{y} - \bar{x}$
76.7, 78.1, 88.5	$s_1 = 28.200$
76.7, 78.1, 97.5	$s_2 = 22.200$
76.7, 78.1, 108.2	$s_3 = 15.067$
76.7, 78.1, 122.2	$s_4 = 5.733$
76.7, 88.5, 97.5	$s_5 = 15.266$
76.7, 88.5, 108.2	$s_6 = 8.133$
76.7, 88.5, 122.2	$s_7 = -1.200$
76.7, 97.5, 108.2	$s_8 = 2.133$
76.7, 97.5, 122.2	$s_9 = -7.200$
76.7, 108.2, 122.2	$s_{10} = -14.334$
78.1, 88.5, 97.5	$s_{11} = 14.334$
78.1, 88.5, 108.2	$s_{12} = 7.200$
78.1, 88.5, 122.2	$s_{13} = -2.133$
78.1, 97.5, 108.2	$s_{14} = 1.200$
78.1, 97.5, 122.2	$s_{15} = -8.133$
78.1, 108.2, 122.2	$s_{16} = -15.266$
88.5, 97.5, 108.2	$s_{17} = -5.733$
88.5, 97.5, 122.2	$s_{18} = -15.067$
88.5, 108.2, 122.2	$s_{19} = -22.200$
97.5, 108.2, 122.2	$s_{20} = -28.200$

2.9 Matrix Representation. The combined sample of these data is naturally represented as a matrix of observations of the form

$$X = \begin{bmatrix} X_{11}^1 & \cdots & X_{1n_1}^1 & \cdots & X_{1n_c}^c \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ X_{p1}^1 & \cdots & X_{pn_1}^1 & \cdots & X_{pn_c}^c \end{bmatrix} \quad (10)$$

where  $X$  is a  $p \times N$  matrix in which the columns are the vector-valued observations; i.e.,

$$X = (X_1^1, \dots, X_{n_1}^1, X_1^2, \dots, X_{n_c}^c). \quad (11)$$

**2.10 Data Analysis.** In the construction of distribution-free procedures, the data are often replaced by their ranks. This transformation may be attractive for a number of reasons, a principal one being the distribution of a test statistic need be established only once. Otherwise, a customized test (such as given in section 2.6) must be constructed for every set of data. The customized test in which the data themselves serve as scores was an idea originally advanced by Fisher (1935)—rank tests and tests based on rank scores are descendants. The development to follow, consistent with Puri and Sen (1993), will use the rank representation, but with the understanding that different score functions remain a viable alternative approach.

**2.11 Rank Matrix Representation.** If the  $N$  observations on the  $i^{\text{th}}$  variate  $X_{ij}^k$ ,  $j = 1, \dots, n_k$ ,  $k = 1, \dots, c$ , are arranged in ascending order, and  $R_{ij}^k$  denotes the rank of  $X_{ij}^k$ , the observation matrix  $X$  gives rise to a corresponding matrix of ranks

$$R = \begin{bmatrix} R_{11}^1 & \dots & R_{1n_1}^1 & \dots & R_{1n_c}^c \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ R_{p1}^1 & \dots & R_{pn_1}^1 & \dots & R_{pn_c}^c \end{bmatrix}. \quad (12)$$

Each row of this matrix is a random permutation of the integers  $1, \dots, N$ , and thus  $R$  is a random matrix that can have  $(N!)^p$  possible realizations.

**2.12 Multivariate Analogue of the Permutation Principle.** Since the  $p$  variates  $X_{ij}^k$ ,  $i = 1, \dots, p$ , are, in general, stochastically dependent, the joint distribution of the elements of  $R$  will depend on the underlying distribution  $F$ , even when the null hypothesis of identical distributions is valid. However, when

$F_1(x) = \dots = F_c(x)$ , the vectors  $X_j^k$ ,  $j = 1, \dots, n_k$ ,  $k = 1, \dots, c$ , are i.i.d. and the joint distribution remains invariant under any permutation of the vectors among themselves. This means the conditional distribution of  $R$  over the set of  $N!$  configurations generated by permutations of the columns of  $R$ , denoted by  $S(R)$ , will be uniform; i.e.,

$$P_0(R = r | S(R)) = 1/N! \quad \forall r \in S(R). \quad (13)$$

This is the multivariate analogue of the permutation principle (section 2.4.1) restricted to rank representation.

2.13 Average Rank Scores. Under  $H_0$ , all the observations  $X_j^k$ ,  $j = 1, \dots, n_k$ ,  $k = 1, \dots, c$ , have the same distribution. Consequently, for each variate  $i$ , the mean of the ranks assigned to the  $k^{\text{th}}$  sample

$$T_i^k = \frac{1}{n_k} \sum_{j=1}^{n_k} R_{ij}^k \quad (14)$$

should be close in value to the overall mean  $\bar{E}_i$ , where

$$\bar{E}_i = \frac{1}{N} \sum_{k=1}^c n_k T_i^k. \quad (15)$$

(The expression for  $\bar{E}_i$  is unnecessarily cumbersome for this application, since  $\bar{E}_i$  is simply the mean of the integers  $1, \dots, N$ ; therefore,  $\bar{E}_i = \frac{(N+1)}{2}$ ,  $i = 1, \dots, p$ . It is written in this form to allow for subsequent inclusion of scores  $a(1), \dots, a(N)$  other than ranks.)

2.14 Rank Order Test Statistic. A test for  $H_0$  based on the contrasts between the mean scores  $T_i^k$  is intuitively appealing. The set of  $p(c-1)$  contrasts  $T_i^k - \bar{E}_i$ ,  $i = 1, \dots, p$ ,  $k = 1, \dots, c$ , should, under  $H_0$ , be numerically small stochastically. For a global assessment of  $H_0$ , a test based on a function of the contrasts that would be sensitive to the numerical largeness

of any contrast seems appropriate. A positive definite quadratic form in  $T_i^k - \bar{E}_i$  will accommodate this. Puri and Sen (1993) advance as a test statistic

$$L_N = \sum_{k=1}^c n_k \left[ (T^k - \bar{E}) V^{-1}(R) (T^k - \bar{E})' \right] \quad (16)$$

where  $T^k = (T_1^k, \dots, T_p^k)$  and  $\bar{E} = (\bar{E}_1, \dots, \bar{E}_p)$ .  $L_N$  is a weighted sum of  $c$  quadratic forms in  $(T^k - \bar{E})$ ,  $k = 1, \dots, c$ , with a common discriminant  $V^{-1}(R)$ .

**2.15 The Discriminant.** The discriminant  $V^{-1}(R)$  is the inverse of the covariance matrix of the variates  $T_i^k$ . There remains to determine the elements of this matrix. Under the null hypothesis of a common distribution,  $T_i^k$  is distributed as the mean of  $n_k$  integers selected at random, without replacement, from the integers  $1, \dots, N$ . The expected value of  $T_i^k$  is then

$$E_0(T_i^k) = \frac{N+1}{2} \quad (\text{Conover 1971}). \quad (17)$$

The notation  $E_0$  denotes expectation under  $H_0$ . To determine the covariance of the variates  $T_i^k, T_{i'}^k$ ,  $i, i' = 1, \dots, p$ , the following result is useful.

• **Lemma.** Let  $X_1, \dots, X_m$  and  $Y_1, \dots, Y_n$  be random samples (not necessarily independent), and let  $\bar{X} = \frac{1}{m} \sum X_i$ ,  $\bar{Y} = \frac{1}{n} \sum Y_j$ , then  $E[\bar{X} \bar{Y}] = E[XY]$ .

• **Proof.** The result is established by the following:

$$\begin{aligned} E[\bar{X} \bar{Y}] &= \frac{1}{m} \sum_{i=1}^m E(X_i \bar{Y}) = E(X \bar{Y}) = \frac{1}{n} \sum_{j=1}^n (X Y_j) \\ &= E[XY]. \end{aligned} \quad (18)$$

This completes the proof.

2.16 The Covariance. The expected value of the product of the sample means is then (from section 2.15.1) the expectation of the product of the variates whose realizations provide the summands. This observation is important, since it permits estimation of the covariance of  $T_i^k, T_{i'}^k, i, i' = 1, \dots, p$ , by the expression

$$\text{Cov}_0(T_i^k, T_{i'}^k) = \frac{1}{N} \sum_{k=1}^c \sum_{j=1}^{n_k} R_{ij}^k R_{i'j}^k - \bar{E}_i \bar{E}_{i'}, \quad (19)$$

analogous to the relation  $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$ .

2.17 The Quadratic Form. The quadratic form  $L_N$  is attractive in that the correlation structure between the variates  $i = 1, \dots, p$ , is taken into account through the covariance matrix  $V(R)$ . Scaling of the variates was simultaneously accomplished by assignment of ranks.

2.18 Application. This methodology will be applied to a communications network simulation validation in section 4.

### 3. TACTICAL NETWORK EXPERIMENT AND SIMULATION DEVELOPMENT

3.1 Background. A controlled laboratory experiment was conducted at the U.S. Army Research Laboratory's (ARL) Command, Control, Communications, and Computers (C4) Research Facility during the summer of 1991 to evaluate the performance of a tactical communications protocol over combat net radios (CNR) (Kaste, Brodeen, and Broome 1992). The approach was to quantify the effects of message arrival rate and message length on the throughput and delay of a small combat radio net. The results provided statistically sound baseline information to be used as input for network simulations, partial guidelines for designing network architectures and communications protocols, and for future experiments on combat radio nets. Eventually a small communications network simulation was developed utilizing the OPNET simulation tool, duplicating the configuration of the aforementioned experiment.

3.2 Experimental Design Factors of Interest. The two factors tested in the experiment were message arrival rate and message length. Four levels of message arrival rate were tested with each of 4 levels of message length (i.e., a full-factorial design), yielding 16 test combinations. The levels for message arrival

rate were 100, 250, 350, and 500 messages per node. The levels for message length were 48, 144, 256, and 352 characters.

- **Design Matrix.** It was decided the shortest reasonable time to test any 1 of the 16 combinations was 1 hr. Since the testing of all 16 combinations required a minimum of 16 hr for a single replication, which realistically could not be completed in 1 day, a randomized incomplete block design was constructed in order that day-to-day variability would not influence the results. The 16 combinations were divided into blocks of size 4, and the 4 blocks were run over a 4-day period. The assignment of the combinations into blocks was based on a confounding scheme. This scheme, in which a different set of three of the 9 degrees of freedom for the interaction term was completely confounded within each replication, assured the effects of message arrival rate and message length, as well as their interaction, on network throughput and network delay could be measured. Three replications of the design matrix were made to ensure the incomplete block design was balanced, thereby facilitating the analysis, although part of the precision of the estimate of the interaction effect was sacrificed (i.e., the relative information available for the interaction term was two-thirds).

**3.3 Experimental and Simulation Configurations.** The experiment consisted of four nodes, each of which was a SUN work station, communicating over a tactical network. Each contained a message driver, providing communications loading, and data collection software to log the sending and receipt of messages and acknowledgments as well as information on queues, as depicted in Figure 1. The nodes were connected to modems to enable communications via radios that could communicate in single-channel (SC) or in frequency-hopping (FH) mode. It was decided to simulate only the SC capability. The modems allowed communication using a specified tactical net-sensing algorithm and communications protocol. To minimize error rates, the radios were placed no more than 3 ft apart and were, therefore, set to low power. Resistor loads were used in place of antennas to avoid interference. The analogous four-node simulation configuration utilizing the OPNET tool is represented in Figure 2. Figure 3 depicts the structure of an individual tactical node. Each node has three processor modules, a queue module that performs the bulk of the channel-access processing, and a pair of radio receiver and transmitter modules.

- **The Server Model.** The four message arrival rates emulated the rate of actual user-generated messages and specific nodes' ability to respond to incoming messages. For the experiment, the arrival rate,  $\lambda$ , represented the number of messages generated during a 1-hr test cell and queued for transmission on the net, not the number of messages actually transmitted during the hour. A message was assumed to

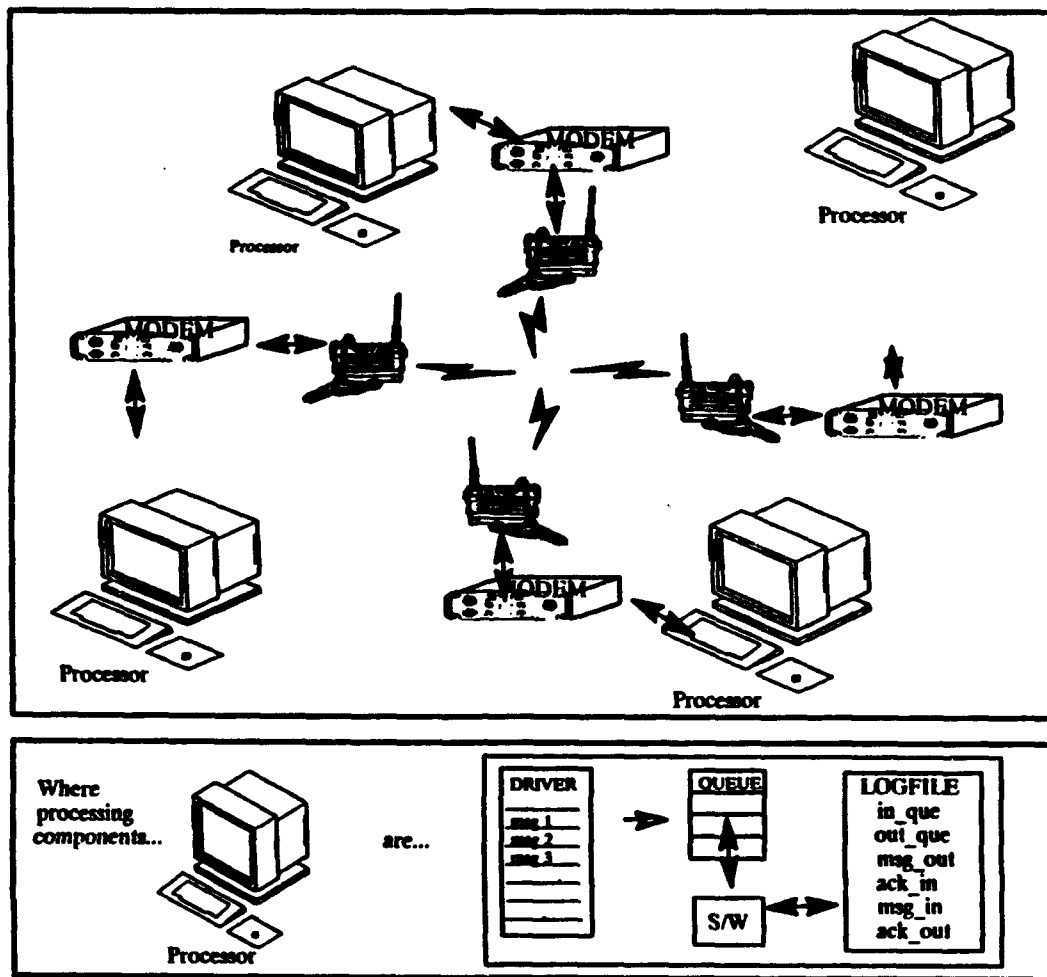


Figure 1. Hardware configuration.

enter network service when it reached the modem, as depicted by the area inside the dashed line in Figure 4. Thus, the server was considered a combination of modem and combat radio network. The queue was the area outside the dashed line. A scenario generator was written to create "messages" of character strings of a specified length and arrival rate over a 1-hr period. The simulation, then, had to accommodate varying message lengths and arrival rates. Once the message was generated, the communications protocol added several layers of information to ensure the message arrived at its destination. This included five error correction/detection bits for each seven-bit character, four synchronization characters, and a preamble to bring the transmitter to full power before the message was sent. Acknowledgments, though shorter in length, were wrapped with similar overhead bits. In the experiment, the numbers of messages generated for transmission each hour by each node were assumed to be mutually independent Poisson-distributed random variables with parameter  $\lambda_i$ . The messages were

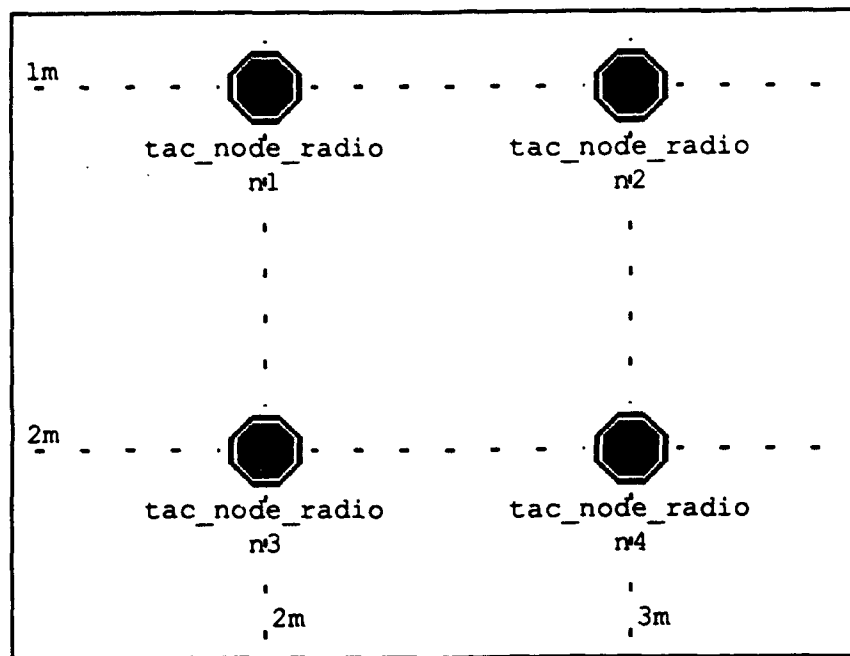


Figure 2. Simulation configuration.

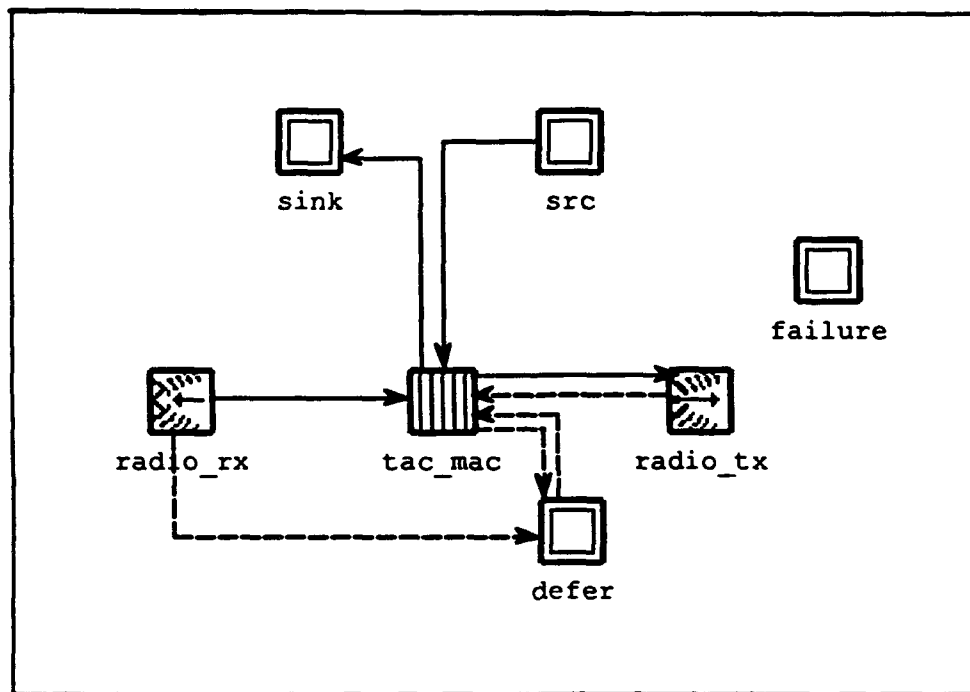


Figure 3. The node model.

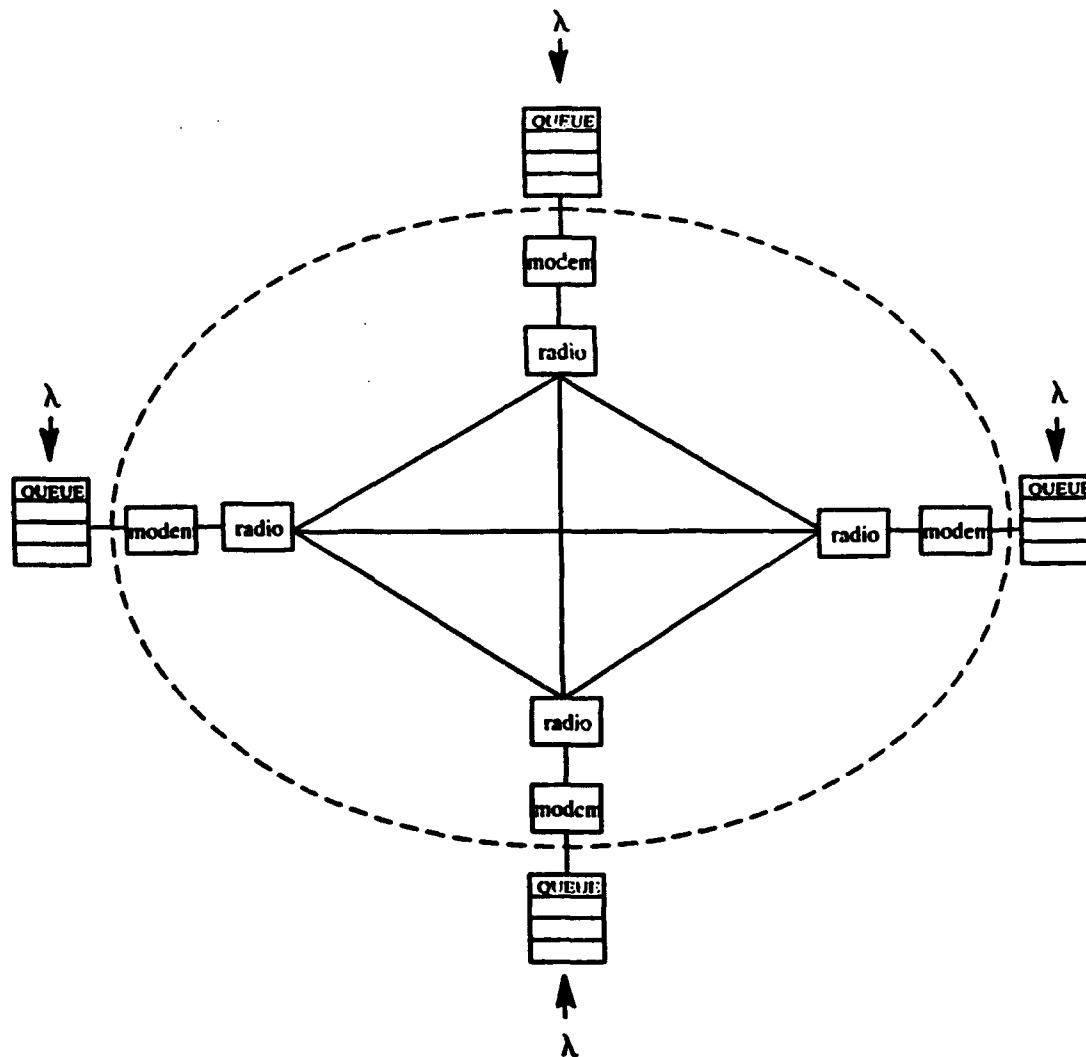


Figure 4. The server model.

equally distributed among the four nodes. For example, if the arrival rate was 2,000 messages/hr, the scenario generator created a file of 500 messages for each node. Each of these details was represented in the simulation. Thus, the simulation represented an actual system using a real and less-than-trivial protocol. In the simulation, the media access control process model (Figure 5) manages the transmission and reception of messages. The tasks are decomposed into three basic functions: encapsulating and queuing outgoing messages, decapsulating and delivering incoming messages, and managing an ongoing transmission.



#### 4. APPLICATION TO NETWORK SIMULATION VALIDATION

4.1 Foreword. Simulations must be subjected to an ongoing validation process before inferences obtained from them about the "real world" can be used with confidence. The permutation test methodology described in section 2 can be utilized to test the validity of a multivariate response simulation with respect to its mean behavior. In this section, the validation test will be applied to a small-scale simulation of a limited bandwidth communications network developed by ARL.

4.2 Reformulation of the Location Problem. Recall that we wish to test the identity of  $c$  ( $\geq 2$ ) multivariate continuous distributions  $F_1, \dots, F_c$ , based on independent random samples from each. Let

$$\mathbf{X}_j^k = (X_{1j}^k, \dots, X_{pj}^k)', \quad j = 1, \dots, n_k, \quad k = 1, \dots, c \quad (20)$$

be such a set of independent vector-valued random samples, where the c.d.f. of  $\mathbf{X}^k$  is denoted by  $F_k(\mathbf{x})$ .

We wish to verify that

$$H_0 : F_1(\mathbf{x}) = \dots = F_c(\mathbf{x}) = F(\mathbf{x}) \quad \forall \mathbf{x} \quad (21)$$

where  $F(\mathbf{x})$  is a common  $p$ -variate c.d.f. We wish to test (see equation 21) against a location parameter-type alternative, that is, vs.

$$H_1 : F_k(\mathbf{x}) = F(\mathbf{x} - \delta_k) \text{ for } k = 1, \dots, c, \text{ and some } \delta_k \neq 0, \quad (22)$$

or equivalently

$$H_0 : \delta_1 = \dots = \delta_c = 0 \quad (23)$$

against the alternative that  $\delta_1, \dots, \delta_c$  are not all equal. Since only a shift in location is being considered, homogeneity of the scale vectors of  $F_1, \dots, F_c$  is assumed.

• Special Case. We consider the special case of comparing two systems (i.e., "real world" and simulated) on the basis of several carefully selected performance measures. We effect this comparison by determining whether  $F_1(\mathbf{x})$  and  $F_2(\mathbf{x})$  differ in location. The data consist of two independent

vector-valued random samples. The  $k^{\text{th}}$  random sample is of size  $n_k$ , where  $k = 1, 2$ . Denote the empirical observations as  $X_j^1$ ,  $j = 1, 2, 3$ ; denote the simulated observations as  $X_j^2$ ,  $j = 1, \dots, 7$ . The total number of observations is  $N = n_1 + n_2 = 3 + 7 = 10$ . There are no missing observations nor tied values to consider.

**4.3 MOP.** Although data for a number of MOP were collected during the experiment, comparisons between empirical and simulation results were limited to network throughput, network delay, and utilization. These were the only measures that could be defined by continuous random variables.

Network throughput is the average number of information bits that were successfully transmitted and acknowledged over a 1-hr test cell. Throughput does not include such overhead as the acknowledgments themselves or, in the event of collisions, message retransmissions. It does, however, include error detection/correction bits and synchronization characters.

Network delay is the average time that passes between a message's arrival at a host's modem until the acknowledgment returns to the host. Messages that were never completely serviced during the running of a test cell were not considered in computing network delay.

Network utilization for a particular time interval is the amount of time spent actually transmitting messages, message retransmissions, or acknowledgments during that interval, divided by the amount of time in the interval. Messages, retransmissions, and acknowledgments include the preamble and other protocol overhead in addition to actual transmission bits.

**4.4 Case Selection.** While 16 combinations of message arrival rate and message length were tested in the 1991 experiment, only 8 were chosen for the validation study. The 8 combinations were not chosen in a purely random fashion as it was desirable to ensure the simulation would be evaluated at the 2 extremes of both parameter ranges (i.e., arrival rate of 400 messages and message length of 48 characters; arrival rate of 2,000 messages and message length of 352 characters). One component of the highest-order interaction was confounded such that the 16 combinations were divided into 2 blocks of

8 units each. The principal block was selected as it contained the two extreme conditions mentioned previously. Given the data were not all collected under the same conditions, each combination was treated as a homogeneous grouping; therefore, each served as an independent case to test the null hypothesis.

- **Observations.** The appropriate empirical observations were taken from each of the three replications performed for the 1991 experiment. The simulation was not run with the scenarios generated for the experimental test cells to ensure the independence of the sample observations. The capability to utilize actual message scenarios as simulation input does, however, afford the developer a useful tool for verification.

**4.5 Testing the Null Hypothesis.** For these data, the number of permutations possible is  $N!$  and the number of distinct values of the  $L_N$  statistic possible is  $\left( \frac{N!}{n_1! n_2! \cdots n_c!} \right) = \frac{N!}{\prod_{k=1}^c n_k!}$ . Since this

validation study deals with small values of  $N$  ( $= 10$ ) and  $p$  ( $= 3$ ), the  $L_N$  statistic may be calculated for all permutations of the observations that lead to distinct values of the statistic; thus, an exact application of the permutation test is possible. We have elected to replace the ranks  $R_{ij}^k$  by a rank score function of the form  $\frac{R_{ij}^k}{(N+1)}$ , thereby reducing  $T_i^k$  to  $(N+1)^{-1}$  times the average rank of the  $k^{\text{th}}$  sample,  $i^{\text{th}}$  variate observations among the combined sample  $i^{\text{th}}$  variate observations. (The motivation for this choice was the univariate case ( $p = 1$ ) for which the statistic  $L_N$  reduces to the Kruskal-Wallis test.) This test was applied to an OPNET communications network simulation in order to validate simulated output with respect to empirical observations. The hypothesis (equation 23) was tested for the MOP outlined in section 4.3. The significance or P-value is the proportion of the  $10!/3!7! = 120$  data permutations providing an equivalent or larger test statistic than that obtained for the reference, or observed, set. Assuming an a priori significance level of 0.05, the null hypothesis was rejected in five of the eight test combinations. The observed test statistic values and the resultant P-levels are summarized in Table 2.

Table 2. Multivariate Multisample Rank Sum Test Results

$H_0: \delta_1 = \dots = \delta_c = 0$				
Input Condition (messages, characters)	Observed Statistic	P-Value	Reject/Fail to Reject	
400, 48	7.184996	0.04167	Reject	
400, 256	9.032534	0.00833	Reject	
1,000, 144	6.970858	0.05833	Fail to Reject	
1,000, 352	9.651814	0.00833	Reject	
1,400, 48	7.826177	0.02500	Reject	
1,400, 256	6.581197	0.10000	Fail to Reject	
2,000, 144	6.734517	0.07500	Fail to Reject	
2,000, 352	9.210527	0.00833	Reject	

4.6 Remarks. We note that the testing procedure and, hence, the conclusion reached, depends on the measure of location shift that is employed. Alternative measures could be employed; however, use of a different estimator for  $\delta$  will likely produce a different conditional rejection region. Since we are free to choose the test statistic, an alternate statistic suggested by Chung and Fraser (1958) was considered.

4.7 Chung and Fraser Test Statistic. The theoretical contributions of Chung and Fraser, while substantial, have generally gone unnoticed. They proposed several randomization tests for the multivariate two-sample problem that were initially developed for the normal-theory two-sample problem for which the Hotelling  $T^2$  test does not exist; however, the tests are also valid in a more general context as nonparametric tests. The approach of Chung and Fraser is to select a statistic suitable for the univariate case, apply it to each of the  $p$  variates, and add the resulting expressions. This approach does not take into account covariances, as is required with the nonparametric counterpart of the Hotelling's  $T^2$  statistic. For measuring shift in location alternatives, a two-sample rank test may be obtained by recording ranks and using the absolute value of the difference in sample means as a test statistic. One of the forms of the Chung and Fraser statistic, and the one considered, is  $\sum_{i=1}^p |\bar{r}_i - \bar{s}_i|$ , appealing in its simplicity (Chung and Fraser 1958).

• **Results.** The observed test statistic values and the resultant P-levels for Chung and Fraser's rank test are summarized in Table 3. Employing the test statistic from section 4.7, the null hypothesis was rejected for six of the eight test conditions. The results were intuitively appealing to the simulation developer who had established validity on the sole basis of visual inspections of the data sets. The conditional rejection region differs from the one established in section 4.5. Indeed, the same conclusion regarding the simulation's validity was established for only three of the eight test conditions. However, Figures 6-8 suggest a correlation structure among the three variates. This was expected from theoretical considerations of the communications network. This structure is not accounted for by the Chung and Fraser statistic.

Table 3. Chung and Fraser Test Statistic Results

$H_0: \delta_1 = \dots = \delta_c = 0$			
Input Condition (messages, characters)	Observed Statistic	P-Value	Reject/Fail to Reject
400, 48	8.809525	0.08333	Fail to Reject
400, 256	7.857143	0.12500	Fail to Reject
1,000, 144	12.619050	0.00833	Reject
1,000, 352	15.000000	0.00833	Reject
1,400, 48	13.571430	0.00833	Reject
1,400, 256	15.000000	0.01667	Reject
2,000, 144	12.619050	0.03333	Reject
2,000, 352	15.000000	0.00833	Reject

## 5. PROJECTION OF CURRENT RESEARCH

**Future Considerations.** As with any well-defined research initiative, additional areas warranting investigation have surfaced, either due to apparent anomalies in the performance of the chosen test statistic or simple curiosity. These areas, however, lie outside the realm of this study, but perhaps should become the focus of future studies. Some possibilities include the following:

Communications data for 400 messages, 48 characters

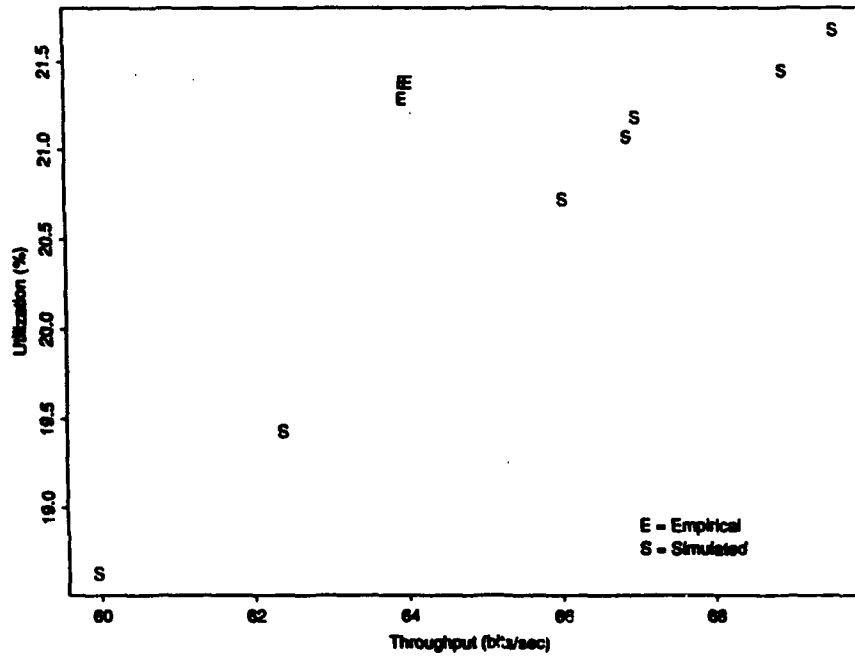


Figure 6. Marginals for a representative data set - throughput vs. utilization.

Communications data for 400 messages, 48 characters

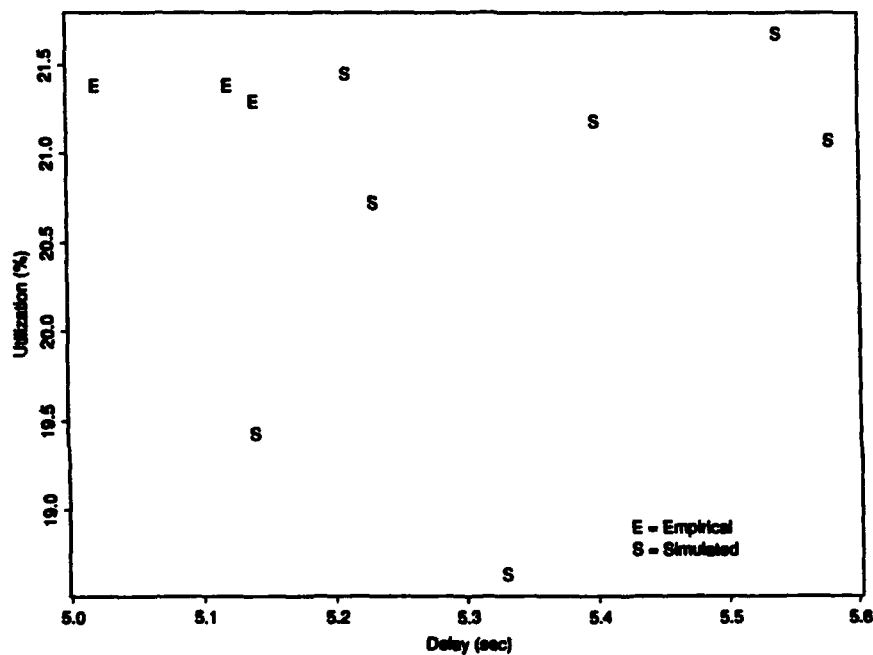


Figure 7. Marginals for a representative data set - delay vs. utilization.

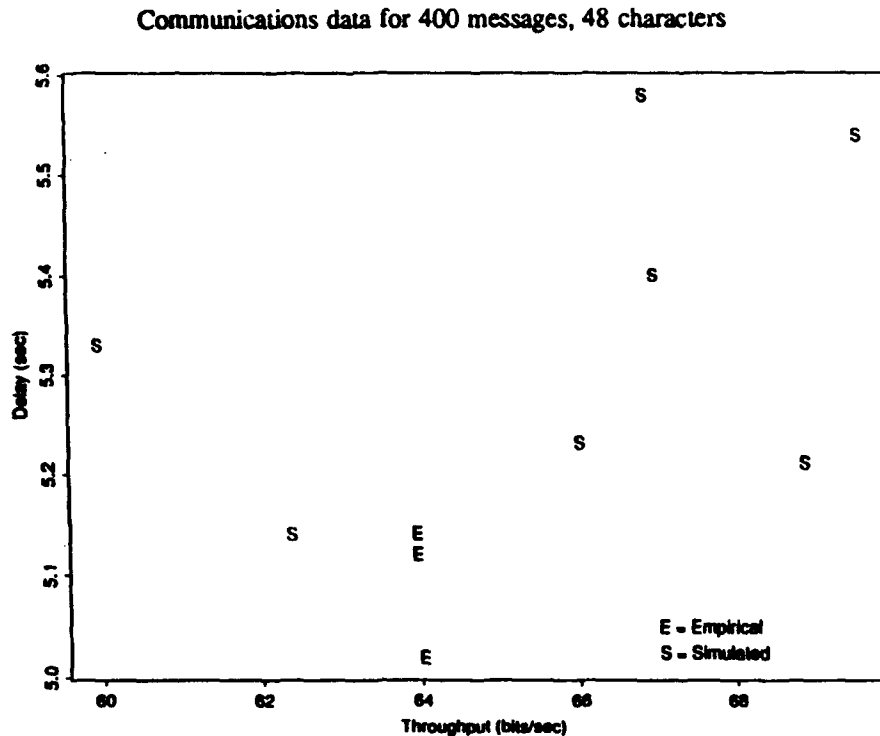


Figure 8. Marginals for a representative data set - throughput vs. delay.

- **Sensitivity of the Multivariate Multisample Test Statistic.** We saw in Tables 2 and 3 different decisions made for the statistic advanced by Puri and Sen (section 2.14) and that of Chung and Frasier (section 4.7). Puri and Sen's statistic  $L_N$ , whose development was considerably more intense, attempts to take into account correlation structure. Chung and Frasier's statistic, which is more direct, does not. Curiously, the decisions associated with the Chung and Frasier statistic may hold more visual appeal (i.e., face validity). Of course, the "curse of dimensionality" is ever present, complicating visual assessment. Reconciliation of the contents of Tables 2 and 3 is a natural consequence of this observation.

- **Combining Independent Tests.** In this validation study, the same null hypothesis was tested for several sets of independent samples, not all necessarily gathered under the same conditions, thereby generating several sets of statistics by which to judge the validity of the communications network simulation. Generally speaking, the military simulation and modeling community prefers a single statistic reflecting the usability of any specific simulation. Given this situation, a possible approach is to combine the various results into a single statistic on which an objective overall judgment can be based. For this, we might begin by considering a technique for combining two-sample tests proposed by van Elteren (1960).

- **Alternative Test Statistics.** Computer-intensive methods may be applied to a variety of hypothesis testing situations. Keeping in mind that we consider these methods as the means by which to generate the probability distribution of some statistic under the "null hypothesis is true assumption," we are free to select and customize a test statistic on the basis of its sensitivity towards an alternative. We may wish to consider other test statistics that measure a difference of location.

- **Generalization of Chung and Fraser Statistic.** In section 4.7.1, we saw that the Chung and Fraser randomization test for the multivariate two-sample problem performed well when considered as an objective counterpart to visual inspection. In fact, it agreed in every case. In their paper, Chung and Fraser (1958) state their methods are easily extended to the k-sample problem. It might be worthwhile to pursue extension of Chung and Fraser's work to the multisample problem.

## 6. SUMMARY

As reliance upon computer simulations to model processes that resist analytical description increases, so does the need to validate the simulations themselves. An impartial approach to simulation validation is through statistical hypothesis testing. In this paper, an application of a nonparametric multivariate procedure to assess the validity of a communications network simulation model, whose intent is to emulate a limited bandwidth combat radio net, is detailed. The procedure, sometimes described as a permutation or randomization test, offers considerable flexibility to the analyst charged with maintaining the fidelity of the modeling effort.

## REFERENCES

- Brodeen, A. E. M., V. A. T. Kaste, and B. D. Broome. "An Experiment to Examine Protocol Performance Over Tactical Networks." Proceedings of the 31st U.S. Army Operations Research Symposium, Fort Lee, VA, pp. C5-C17, November 1992.
- Chung, J. H., and D. A. S. Fraser. "Randomization Tests for a Multivariate Two-Sample Problem." Journal of the American Statistical Association, vol. 53, pp. 729-735, 1958.
- Conover, W. J. Practical Nonparametric Statistics. New York: John Wiley & Sons, Inc., 1971.
- Edgington, E. S. Randomization Tests. New York: Marcel Dekker, Inc., 1987.
- Fisher, R. A. Design of Experiments. Edinburgh: Oliver and Boyd, 1935.
- Kaste, V. A. T., A. E. M. Brodeen, and B. D. Broome. "An Experiment to Examine Protocol Performance Over Combat Net Radios." BRL-MR-3978, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, MD, June 1992.
- Noreen, E. W. Computer-Intensive Methods for Testing Hypotheses: An Introduction. New York: John Wiley & Sons, Inc., 1989.
- Puri, M. L., and P. K. Sen. Nonparametric Methods in Multivariate Analysis. Malabar, FL: Krieger Publishing Company, 1993.
- Randles, R. H., and D. A. Wolfe. Introduction to the Theory of Nonparametric Statistics. New York: John Wiley & Sons, Inc., 1979.
- Sargent, R. G. "Validation and Verification of Simulation Models." Proceedings of the 1992 Winter Simulation Conference, Arlington, VA, pp. 104-114, 1992.
- Taylor, M. S. "A Randomization Approach to an Analysis of Ballistic Data." Proceedings of the 37th Conference on the Design of Experiments in Army Research, Development, and Testing, Vicksburg, MS, pp. 47-52, May 1992.
- van Elteren, P. "On the Combination of Independent Two-Sample Tests of Wilcoxon." Bulletin de l'Institut International de Statistique, vol. 37, 1960.

INTENTIONALLY LEFT BLANK.

## BIBLIOGRAPHY

- Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley & Sons, Inc., 1958.
- Balci, O., and R. G. Sargent. "Validation of Multivariate Response Models Using Hotelling's Two-Sample T Test." Simulation, pp. 185-192, December 1982.
- Boyett, J. M., and J. J. Shuster. "Nonparametric One-Sided Tests in Multivariate Analysis with Medical Applications." Journal of the American Statistical Association, vol. 72, pp. 665-668, 1977.
- Garratt, M. "Statistical Validation of Simulation Models." Proceedings of the 1974 Summer Simulation Conference, Houston, TX, pp. 915-926, 1974.
- Good, P. Permutation Tests. New York: Springer-Verlag, 1994.
- Green, P., Jr. Computer Network Architectures and Protocols. Edited by F. A. Tobagi. New York: Plenum Press, 1982.
- Kleinrock, L. Communications Nets - Stochastic Message Flow and Delay. New York: McGraw-Hill Book Co., 1964.
- Lehmann, E. L. Nonparametrics: Statistical Methods Based on Ranks. New York: Holden-Day, Inc., 1975.
- Magnavox Government and Industrial Electronics Company. "AFATDS Performance Modeling System Level for Multiservice Advanced Field Artillery Tactical Data System." Contract No. DAAB07-89-G-E424-0001, CDRL Sequence No. NA, Ft. Wayne, IN, 1991.
- Morrison, D. F. Multivariate Statistical Methods. New York: McGraw-Hill Book Company, 1976.
- Reynolds, M. R., H. E. Burkhardt, and R. F. Daniels. "Procedures for Statistical Validation of Stochastic Simulation Models." Forest Science, vol. 27, no. 2, pp. 349-364, 1981.
- Reynolds, M. R., and M. L. Deaton. "Comparisons of Some Tests for Validation of Stochastic Simulation Models." Communications in Statistics - Simulation and Computation, vol. 11, no. 6, pp. 769-799, 1982.
- Sen, P. K. Electronic-mail communication with author. University of North Carolina, Chapel Hill, NC, 30 November 1993.
- Urquhart, S. N. "On the Permutation Distribution of a Multivariate Test Statistic." Ph.D. dissertation, Colorado State University, Fort Collins, CO, 1965.

INTENTIONALLY LEFT BLANK.

<u>No. of Copies</u>	<u>Organization</u>
2	Administrator Defense Technical Info Center ATTN: DTIC-DDA Cameron Station Alexandria, VA 22304-6145
1	Commander U.S. Army Materiel Command ATTN: AMCAM 5001 Eisenhower Ave. Alexandria, VA 22333-0001
1	Director U.S. Army Research Laboratory ATTN: AMSRL-OP-SD-TA, Records Management 2800 Powder Mill Rd. Adelphi, MD 20783-1145
3	Director U.S. Army Research Laboratory ATTN: AMSRL-OP-SD-TL, Technical Library 2800 Powder Mill Rd. Adelphi, MD 20783-1145
1	Director U.S. Army Research Laboratory ATTN: AMSRL-OP-SD-TP, Technical Publishing Branch 2800 Powder Mill Rd. Adelphi, MD 20783-1145
2	Commander U.S. Army Armament Research, Development, and Engineering Center ATTN: SMCAR-TDC Picatinny Arsenal, NJ 07806-5000
1	Director Benet Weapons Laboratory U.S. Army Armament Research, Development, and Engineering Center ATTN: SMCAR-CCB-TL Watervliet, NY 12189-4050
1	Director U.S. Army Advanced Systems Research and Analysis Office (ATCOM) ATTN: AMSAT-R-NR, M/S 219-1 Ames Research Center Moffett Field, CA 94035-1000

<u>No. of Copies</u>	<u>Organization</u>
1	Commander U.S. Army Missile Command ATTN: AMSMI-RD-CS-R (DOC) Redstone Arsenal, AL 35898-5010
1	Commander U.S. Army Tank-Automotive Command ATTN: AMSTA-JSK (Armor Eng. Br.) Warren, MI 48397-5000
1	Director U.S. Army TRADOC Analysis Command ATTN: ATRC-WSR White Sands Missile Range, NM 88002-5502
1	Commandant U.S. Army Infantry School ATTN: ATSH-WCB-O Fort Benning, GA 31905-5000
<u>Aberdeen Proving Ground</u>	
2	Dir, USAMSAA ATTN: AMXSY-D AMXSY-MP, H. Cohen
1	Cdr, USATECOM ATTN: AMSTE-TC
1	Dir, USAERDEC ATTN: SCBRD-RT
1	Cdr, USACBDCOM ATTN: AMSCB-CII
1	Dir, USARL ATTN: AMSRL-SL-I
5	Dir, USARL ATTN: AMSRL-OP-AP-L

No. of  
Copies Organization

Aberdeen Proving Ground

17 Dir, USARL  
ATTN: AMSRL-SL-BV,  
W. Winner  
W. Baker  
AMSRL-WT-PB, D. Webb  
AMSRL-CI,  
William H. Mermagen  
B. Bodt  
M. Taylor (3 cp)  
AMSRL-CI-CC,  
R. Kaste  
A. Brodeen (7 cp)  
AMSRL-CI-CA, B. Broome

## USER EVALUATION SHEET/CHANGE OF ADDRESS

This Laboratory undertakes a continuing effort to improve the quality of the reports it publishes. Your comments/answers to the items/questions below will aid us in our efforts.

1. ARL Report Number ARL-TR-592 Date of Report October 1994
2. Date Report Received \_\_\_\_\_
3. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which the report will be used.) \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
4. Specifically, how is the report being used? (Information source, design data, procedure, source of ideas, etc.) \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
5. Has the information in this report led to any quantitative savings as far as man-hours or dollars saved, operating costs avoided, or efficiencies achieved, etc? If so, please elaborate. \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
6. General Comments. What do you think should be changed to improve future reports? (Indicate changes to organization, technical content, format, etc.) \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

### CURRENT ADDRESS

\_\_\_\_\_  
Organization

\_\_\_\_\_  
Name

\_\_\_\_\_  
Street or P.O. Box No.

\_\_\_\_\_  
City, State, Zip Code

7. If indicating a Change of Address or Address Correction, please provide the Current or Correct address above and the Old or Incorrect address below.

### OLD ADDRESS

\_\_\_\_\_  
Organization

\_\_\_\_\_  
Name

\_\_\_\_\_  
Street or P.O. Box No.

\_\_\_\_\_  
City, State, Zip Code

(Remove this sheet, fold as indicated, tape closed, and mail.)  
(DO NOT STAPLE)

---

**DEPARTMENT OF THE ARMY**

**OFFICIAL BUSINESS**



**NO POSTAGE  
NECESSARY  
IF MAILED  
IN THE  
UNITED STATES**

**BUSINESS REPLY MAIL**  
**FIRST CLASS PERMIT NO 0001, APG, MD**

Postage will be paid by addressee

**Director**  
**U.S. Army Research Laboratory**  
**ATTN: AMSRL-OP-AP-L**  
**Aberdeen Proving Ground, MD 21005-5066**

